Network Growth with Preferential Attachment for High Indegree and Low Outdegree

Volkan Sevim[1*] and Per Arne Rikvold[1,2†]
[1] *School of Computational Science, Center for Materials Research and Technology, and Department of Physics, Florida State University, Tallahassee, FL 32306-4120, USA*
[2] *National High Magnetic Field Laboratory, Tallahassee, FL 32310-3706, USA*
(Dated: March 17, 2008)

We study the growth of a directed transportation network, such as a food web, in which links carry resources. We propose a growth process in which new nodes (or species) preferentially attach to existing nodes with high indegree (in food-web language, number of prey) and low outdegree (or number of predators). This scheme, which we call *inverse preferential attachment*, is intended to maximize the amount of resources available to each new node. We show that the outdegree (predator) distribution decays at least exponentially fast for large outdegree and is continuously tunable between an exponential distribution and a delta function. The indegree (prey) distribution is poissonian in the large-network limit.

I. INTRODUCTION

Directed networks that transport a resource, such as energy, from one or several sources to a large number of consumers are important in many areas of science [1]. Among such networks, food webs provide an example of great interest, both from a purely scientific point of view, and because of their importance for nature-conservation efforts [2]. Although knowledge is rapidly accumulating about the *structure* of food webs [2, 3, 4, 5, 6, 7], and static models have been developed to describe some of their statistical properties [3, 4, 5, 8], we are as yet only beginning to develop an understanding of the processes by which such networks are formed and evolve under the influence of speciation, invasion, and extinction of interacting species [9, 10, 11, 12, 13, 14].

An important aspect of the network structure of food webs is that they have degree distributions that generally decay quite fast with increasing degree – in most cases at least exponentially [4, 5, 8, 13]. This is in sharp contrast to the class of networks known as scale-free, which have power-law degree distributions [15, 16]. While there has been a veritable explosion of research on scale-free networks, there has been no similar surge of interest in networks with rapidly convergent degree distributions. Most food webs, and some (but not all [17]) transportation networks, such as the North American power grid [18], and the European railway network [19] belong to this class. Much work remains to be done before a comprehensive understanding of the mechanisms by which such networks evolve is reached.

As a step toward the development of such an understanding, we here propose a network growth scheme that produces a poissonian indegree distribution (in food-web language: prey distribution) and an outdegree (predator) distribution that is continuously tunable between an exponential distribution and a delta function. We note that these degree distributions do not agree with current food-web theory. In particular, the indegree distribution produced by the niche model [8] has an exponential tail [4]. It has been claimed that models with an exponentially decaying probability of preying on a given fraction of species with lower or equal niche values are capable of producing food webs that are structurally in agreement with empirical data [6]. However, it is not clear why this condition is necessary or why schemes that invoke no physical mechanisms (as in the niche model) are able to produce such webs. Therefore, other plausible schemes should also be explored.

Our model employs a scheme in which new nodes (species) attach to existing nodes with a preference for nodes i with high indegree k_i' and low outdegree k_i. In food-web terms, this corresponds to a prospective predator choosing prey that have a large number of resources (represented by the large indegree), while the competition from previously established predators should be as small as possible (low outdegree). Among the influences on the network growth process mentioned above (speciation, invasion, and extinction), we have thus chosen to focus on invasion and/or speciation. By ignoring extinction, we essentially model the early phase of steady network growth. The proposed growth process corresponds to a probability of attachment,

$$\Pi(k_i', k_i) \propto (k_i'/k_i)^\gamma \tag{1}$$

*Electronic address: sevim@scs.fsu.edu
†Electronic address: rikvold@scs.fsu.edu

with $\gamma \geq 0$. This attachment scheme is the direct opposite of the "rich get richer" scheme of preferential attachment with an attachment probability proportional to the total degree $l_i = k'_i + k_i$, which is known to produce scale-free networks and has been studied in a myriad of variations over the last decade [15, 16]. To emphasize this difference, we shall call the scheme proposed here *inverse preferential attachment*. [We note that nonlinear forms of the "rich get richer" scheme with a probability of attachment proportional to l_i^α with $\alpha > 0$ have also been studied. However, except for $\alpha = 1$, no $\alpha > 0$ leads to a power-law degree distribution [20].]

In a previous paper [21], we studied a simplified version of the scheme proposed here, in which a new node makes a constant number of incoming links ($k'_i = \text{const.}$), and $\gamma = 1$. In this simplified version, the probability of attachment, $\Pi(k_i) \propto 1/k_i$, depends only on the outdegree (the number of predators). We calculated the outdegree distribution for this simplified model both analytically and by Monte Carlo simulations. It is given by the self-consistent equation

$$n_k^* = (k+1)(m/z_m^*)^k \frac{\Gamma\big(1 + m/z_m^*\big)}{\Gamma\big(k + 2 + m/z_m^*\big)} , \tag{2}$$

where

$$z_m^* = \sum_{j=0}^{\infty} \frac{n_j^*}{j+1} , \tag{3}$$

and $\Gamma(x)$ represents the Gamma function.

II. MODEL AND RESULTS

In the present paper, we investigate the general form of the attachment probability presented in Eq. (1). With this form we relax both of the restrictions of the simplified model: we do not fix the indegree (number of prey) for the new nodes, so that each can make a different number of links, and we also vary the exponent γ. This makes full analytical treatment much harder, and the results for the outdegree distribution presented here are therefore only numerical. The indegree distribution, however, is found analytically to be a poissonian in the large-network limit.

The generalized attachment process proceeds as follows. We start the growth process with N_0 nodes and assign each initial node an indegree $m \leq N_0$. These nodes act as source nodes as they are not connected to any other node at the beginning. Actually, the attributes of the initial nodes have no significance for the statistics because the final size of the network, $N_{\max} + N_0$, is much larger than N_0. In each time step, we add a new isolated node. Then, we give the new node m chances to establish a link to an existing node with probability

$$\Pi(k'_i, k_i) = \frac{1}{Z} \left(\frac{k'_i}{k_i + 1} \right)^{\gamma} , \tag{4}$$

where

$$Z = \sum_{i=1}^{N} \left(\frac{k'_i}{k_i + 1} \right)^{\gamma} \tag{5}$$

with $\gamma \geq 0$. Here, N denotes the number of existing nodes at that time step. We use $k_i + 1$ in the denominator to prevent a divergence for $k_i = 0$. Multiple links between two nodes are not allowed. The direction of a link is from the old node to the new one.

We implement the growth process in Monte Carlo simulations as follows. We seed the system with N_0 source nodes, each with indegree m, and introduce a new node in each Monte Carlo step. To create links between the new node and the existing ones, we pick existing nodes, i, one by one and calculate the probabilities of attachment, $\Pi(k'_i, k_i)$. Then, we generate a random number, r, and attach the new node to node i if $r < \Pi$. We repeat this procedure until all existing nodes in the network are tested, i.e., till the sweep is completed. Since $\sum_i \Pi(k'_i, k_i) = 1$, a new node makes on average one connection per sweep. We sweep the whole network m times, so that $\langle k'_i \rangle = \langle k_i \rangle = m$. The new node is kept in the system, even if it does not acquire any links. However, a node with $k' = 0$ stays isolated throughout the growth since the probability of attachment to it is zero. We stop the growth when the network size N reaches $N_{\max} + N_0$ nodes with $N_{\max} = 10^5$. We average over fifteen independent runs for each value of m and γ.

We first tested the case of $\gamma = 1$ to compare the outdegree distribution of the full k'/k model to the outdegree distribution of our simplified $1/k$ model, Eqs. (2-3), which, for large k, decays like $k\mu^k/\Gamma(k)$, where μ is a constant. As seen in Fig. 1, the outdegree distribution for the k'/k model also decays faster than exponentially for large k.

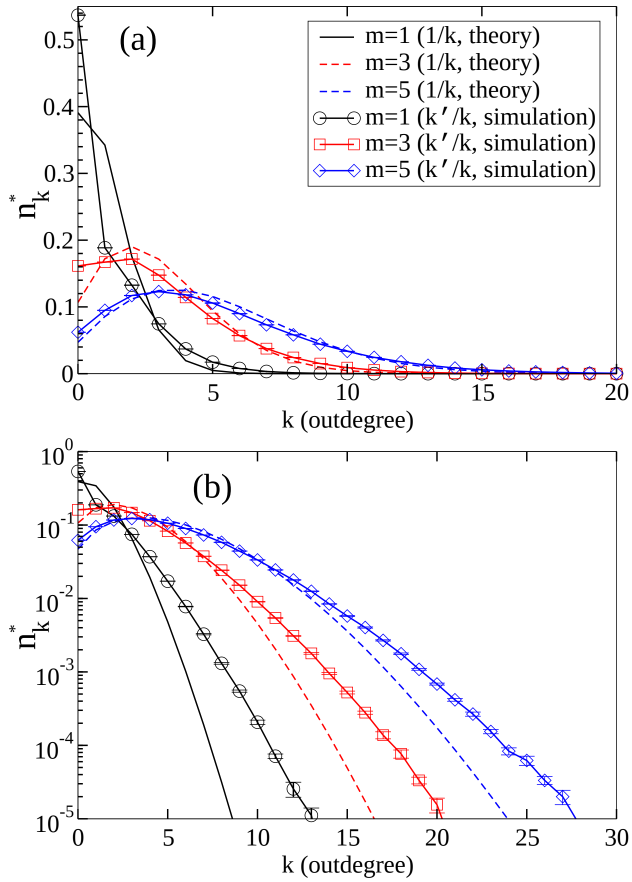

FIG. 1: Outdegree distributions for $\gamma = 1$ and $m = 1$, 3, and 5 with $N_0 = 10$ shown on linear (a) and log-linear (b) scales. The simulations were stopped when the network size reached $N_0 + 10^5$ nodes. Each curve (with symbols) represents an average over fifteen runs. The curves without symbols are the theoretical outdegree distributions for our simplified model with $\Pi(k) \propto 1/k$. Both the k'/k and $1/k$ models yield the same distribution for $m \gg 1$. As k is a discrete variable, lines connecting the symbols are merely guides to the eye.

However, the dependence on the variable indegree leads to a broadening of the outdegree distribution: decreased probabilities for $k \approx m$, and compensating increased probabilities for $k \gg m$ and $k \ll m$. In the limit of large m, the central part of the outdegree distribution of the k'/k model approaches that of the $1/k$ model.

The outdegree distribution of the general model also varies with γ. Higher values of γ sharpen the peak of the distribution around the mean outdegree, m, as it increases the tendency of the new nodes to prefer existing nodes with a higher value of k'/k (Fig. 2). In the limit $\gamma \to \infty$ one should obtain a delta function at $k = m$. Similarly, lower values of γ relax the constraint and flatten the outdegree distribution. The limiting case, $\gamma = 0$, corresponds to growth without preferential attachment, which yields an exponential outdegree distribution of mean m [16, 21]. The lines connecting the symbols are guides to the eye.

In contrast to the outdegree distribution, the indegree distribution of the generalized model in the $N \gg m, k'$ limit can be described analytically and is extremely well approximated by a Poisson distribution with mean m, independent of γ (Fig. 3). This can be shown as follows. Each new node makes one link per sweep on average. When $N \gg m$, each existing node has a probability equal to $1/N$ of acquiring a new link per sweep on average, independent of the history of the network. Therefore, the probability of acquiring k' links (after m sweeps) for the node added at time t, when the total number of nodes is N, is a binomial,

$$P_{k'}(N) = \binom{N}{k'} p^{k'} (1-p)^{N-k'} \tag{6}$$

with $p = m/N$. Thus, the final indegree distribution is an average (ignoring the N_0 initial nodes),

$$n^*_{k'} = \frac{1}{N_{\max}} \sum_{N=N_0}^{N_0+N_{\max}} P_{k'}(N) \; . \tag{7}$$

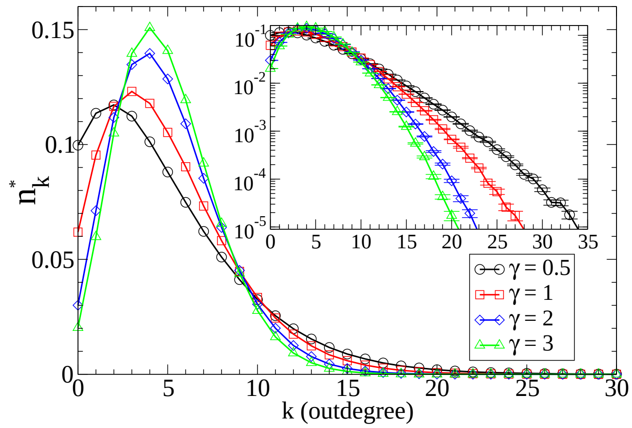

FIG. 2: Outdegree distributions for $\gamma = 0.5$, 1, 2, and 3, with network size $N_0 + 10^5$ nodes, and $m = 5$. Each curve is averaged over fifteen runs. Inset: The same distributions shown on a log-linear scale. The lines connecting the symbols are guides to the eye.

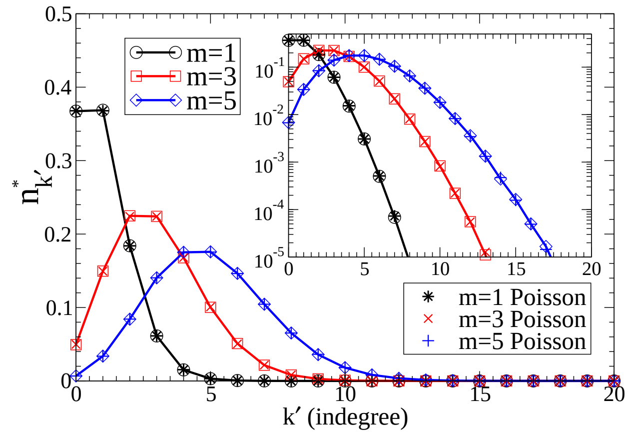

FIG. 3: Indegree distributions for $\gamma = 1$ and $m = 1$, 3, and 5 with $N_0 = 10$. The simulations were stopped when the network size reached $N_0 + 10^5$ nodes. Each curve represents an average over fifteen runs. The error bars are smaller than the symbol sizes. Inset: The same distributions shown on a log-linear scale. The symbols $*$, $\times$, and $+$ show the Poisson distribution, Eq. (8), for $m = 1$, 3, and 5, respectively. The lines connecting the symbols are guides to the eye. See text for details.

In general, this cannot be calculated exactly. However, for $N \gg m, k'$, $P_{k'}(N)$ can be approximated by a poissonian of mean m [22],

$$P_{k'} = \frac{m^{k'} \exp(-m)}{k'!} \ , \tag{8}$$

which is independent of N. The convergence with N to this result is fast, so that for $N_{\max} \gg N_0, m$, the sum in Eq. (7) is dominated by the N-independent terms. As a result,

$$n^*_{k'} \approx \frac{m^{k'} \exp(-m)}{k'!} \tag{9}$$

is an excellent approximation, as shown in Fig. 3. Computer simulations confirm the γ-independent form of the indegree distribution, as shown in Fig. 4.

III. DISCUSSION

The mechanism that we propose and study in this paper, growth by preference for high indegree (number of prey) and low outdegree (number of predators), is intended to simulate the early stages of the development of food webs

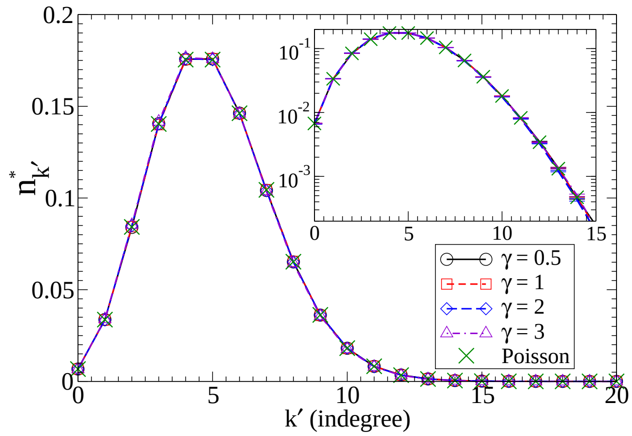

FIG. 4: Indegree distributions for $\gamma = 0.5$, 1, 2, and 3 with $m = 5$, and network size $N_0 + 10^5$ nodes. Each curve is averaged over fifteen runs. The distributions are identical and they all practically overlap with the Poisson distribution, Eq. (8), with $m = 5$. Inset: The same distributions shown on a log-linear scale. The lines connecting the symbols are guides to the eye.

and other transportation networks. To distinguish it from the more commonly studied "rich get richer" schemes that produce scale-free networks, we call it *inverse preferential attachment*.

The outdegree distribution obtained using the generalized form of the probability of attachment, $\Pi(k'_i, k_i) \propto (k'_i/k_i)^\gamma$, (with $\gamma = 1$) is broader than the one obtained from the simplified form, $\Pi(k_i) \propto 1/k_i$. However, both decay faster than exponentially for large k_i. The shape of the outdegree distribution is continuously tunable by γ, from a exponential distribution for $\gamma = 0$ to a delta function for $\gamma \to \infty$. The indegree distribution does *not* depend on γ, but only on m, the mean number of links per node. In the limit $N \gg k, m$, the indegree distribution is a poissonian.

Some features of the networks produced by this model, like the outdegree distribution, which decays faster than exponentially, resemble those of some empirical and model food webs [4, 5, 8, 12, 13]. In this, they differ sharply from the scale-free networks generated by the conventional "rich get richer" preferential-attachment schemes [15, 16]. However, differences from real food webs remain, such as the correlation between the in- and outdegrees of a node. Our model produces webs with a positive in-outdegree correlation, whereas most empirical and model webs have a negative correlation [6, 13]. This may be due to the unrestrained growth of our networks, which would require an extinction process to achieve a steady state [12]. Also, this growth scheme is not designed to produce loops. We intend to include such features in future versions of the model, thus enabling modeling of mature, steady-state networks.

Acknowledgments

This research was supported by U.S. National Science Foundation Grant Nos. DMR-0240078 and DMR-0444051, and by Florida State University through the Department of Physics, the School of Computational Science, the Center for Materials Research and Technology, and the National High Magnetic Field Laboratory.

[1] J. R. Banavar, A. Maritan, A. Rinaldo, Size and form in efficient transportation networks, Nature (London) 399 (1999) 130 – 132.

[2] N. Rooney, K. McCann, G. Gellner, J. C. Moore, Structural asymmetry and the stability of diverse food webs, Nature (London) 442 (2006) 265–269.

[3] J. E. Cohen, F. Briand, C. E. Newman, Community Food Webs, Biomathematics Vol. 20, Springer-Verlag, Berlin, 1990.

[4] J. Camacho, R. Guimerà, L. A. Nunes Amaral, Robust patterns in food web structure, Phys. Rev. Lett. 88 (2002) 228102.

[5] J. Camacho, R. Guimerà, L. A. Nunes Amaral, Analytical solution of a model for complex food webs, Phys. Rev. E 65 (2002) 030901(R).

[6] D. Stouffer, J. Camacho, R. Guimerà, C. A. Ng, L. A. Nunes Amaral, Quantitative patterns in the food-web structure, Ecology 86 (2005) 1301–1311.

[7] D. B. Stouffer, J. Camacho, R. Guimerà, L. A. Nunes Amaral, A robust measure of food web intervality, Proc. Natl. Acad. Sci. USA 103 (2006) 19015–19020.

[8] R. J. Williams, N. D. Martinez, Simple rules yield complex food webs, Nature (London) 404 (2000) 180 – 183.

[9] B. Drossel, P. G. Higgs, A. J. McKane, The influence of predator-prey population dynamics on the long-term evolution of food web structure, J. Theor. Biol. 208 (2001) 91–107.
[10] B. Drossel, A. McKane, C. Quince, The impact of non-linear functional responses on the long-term evolution of food web structure, J. Theor. Biol. 229 (2004) 539–548.
[11] C. Quince, P. G. Higgs, A. J. McKane, Topological structure and interaction strengths in model food webs, Ecol. Modell. 187 (2005) 389–412.
[12] A. G. Rossberg, H. Matsuda, T. Amemiya, K. Itoh, Some properties of the speciation model for food-web structure - mechanisms for degree distributions and intervality, J. Theor. Biol. 238 (2006) 401–415.
[13] P. A. Rikvold, V. Sevim, An individual-based predator-prey model for biological coevolution: Fluctiations, stability and community structure, Phys. Rev. E 75 (2007) 051920.
[14] P. A. Rikvold, Self-optimization, community stability, and fluctuations in two individual-based models of biological coevolution, J. Math. Biol. 55 (2007) 653–677.
[15] A. L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509 – 512.
[16] R. Albert, A. L. Barabasi, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47 – 97.
[17] R. Guimerà, L. A. N. Amaral, Modeling the world-wide airport network, Eur. Phys. J. B 38 (2004) 381 – 385.
[18] R. Albert, I. Albert, G. L. Nakarado, Structural vulnerability of the North American power grid, Phys. Rev. E 69 (2004) 025103(R).
[19] M. Kurant, P. Thiran, Layered complex networks, Phys. Rev. Lett. 96 (2006) 138701.
[20] P. L. Krapivsky, S. Redner, F. Leyvraz, Connectivity of growing random networks, Phys. Rev. Lett. 85 (2000) 4629 – 4632.
[21] V. Sevim, P. A. Rikvold, Effects of preference for attachment to low-degree nodes on the degree distributions of a growing directed network and a simple food-web model, Phys. Rev. E 73 (2006) 056115.
[22] V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley, New York, 1976.